FISHLIT, A REVIEW OF NISC SOUTH AFRICA'S AQUATIC SCIENCES DATABASE

Amanda-Sue McPhail
NISC South Africa
22 Somerset Street
P.O. Box 377
Grahamstown, South Africa, 6139
fishlit@ru.ac.za

Margaret Crampton
NISC South Africa
22 Somerset Street
P.O. Box 377
Grahamstown, South Africa, 6139
nisc@ru.ac.za

ABSTRACT: FISHLIT, a bibliographic database on fish, fisheries and aquaculture produced by NISC South Africa, has been in existence since 1985 and is currently available on CD-ROM, being included on the NISC publications Aquatic Biology, Aquaculture and Fisheries Resources, and Fish and Fisheries Worldwide, and the Internet. An historical account of the design and development of FISHLIT, including a description of the data structure, database size and growth, is provided. A breakdown of source material, for example scientific articles, reports, popular articles and books, is given, as well as the grey literature indexed, a notable strength of this database. A bibliometric analysis is made of the two anthology CD-ROM titles on which FISHLIT is published in terms of subject coverage and numbers of records. Topic coverage is graphically presented and the application of FISHLIT to areas of aquatic science, such as ichthyology, aquaculture and fisheries management, is discussed. Comparisons are drawn between FISHLIT and other currently available aquatic science databases.

KEYWORDS: information technology, aquatic science, bibliographic database, fish, aquaculture, fisheries, data structure, database production, data collection

Background

NISC, founded in 1988, is an electronic publishing company with headquarters in Baltimore, Maryland in the United States. There are six NISC branches world-wide, including the United States, South Africa, Singapore, India, Chile and Mexico, and NISC currently publishes more than one hundred bibliographic and full-text databases on 53

CD-ROM titles. In addition to publishing databases, NISC compiles several databases and markets all NISC products worldwide. NISC South Africa in particular markets all NISC products throughout Africa and compiles the FISHLIT database.

FISHLIT is a bibliographic database on fish, fisheries and aquaculture produced by NISC South Africa in co-operation with the JLB Smith Institute of Ichthyology in Grahamstown, and the Oceanographic Research Institute in Durban, South Africa. Initiated in 1985 to computerise the information function of the world renowned JLB Smith Institute of Ichthyology Library, FISHLIT has grown into an internationally recognised database used by a wide variety of aquatic scientists. FISHLIT is now published on the two NISC discs, Fish and Fisheries Worldwide (FFW), and Aquatic Biology, Aquaculture and Fisheries Resources (ABAFR).

NISC disc features

Discs published by NISC cover a wide variety of topics and a particular feature of NISC CD-ROMs is the anthology disc. These are single discs on which suites of closely related databases are integrated, offering a platform for the publication of large and small databases, integrating diverse resources, creating new, unique database combinations, and significantly increasing the size of the searchable information storehouse. A valuable advantage of anthology publishing is that of composite records. Databases are not mutually exclusive and in many instances a number of database producers may enter the same record in each of their respective databases. Special duplicate detection software designed for NISC discs resolves this problem by combining all the information that is shared between common records, as well as any unique data, to create a kind of super record or composite record. The sum of several records is more valuable than the single record alone as keywords are combined and abstracts added, enhancing the record's searchability.

FISHLIT: A brief historical account

In 1984 a decision was taken to computerise the holdings of the well-known JLB Smith Institute of Ichthyology Library, formerly known as the Fish Library and Information Centre, in Grahamstown, South Africa (Crampton 1988). This decision led, in 1985, to the design and development of the FISHLIT database by Margaret Crampton, then the librarian at the JLB Smith Institute. Initially produced by the JLB Smith Institute library in co-operation with the South African Water Information Centre (Crampton 1988), FISHLIT was later, in 1990, acquired by NISC USA to publish on their Fish and Fisheries Worldwide disc. Realising a need for indigenous information and local service NISC USA, headed by Fred Durr, approached Margaret to run a NISC for Africa and in May 1995 NISC South Africa was founded. A lull in the production of FISHLIT followed as attention focussed on various other aspects of NISC, such as developing discs and licensing databases. Subsequently, in the mid-nineties a permanent staff member was employed to build FISHLIT along with several students who assisted part-time with the indexing. A second permanent staff member has since been employed and the part-time staff number increased to eight. Strong links are still maintained with the JLB Smith

Institute library that provides many of the source materials and assists with document delivery.

Database design

FISHLIT is a bibliographic database, each record representing a publication and made up of a number of searchable fields. These include the author, or authors, of the publication. Author information is stored in a sub-database, as is much of the additional information included in FISHLIT, and in many cases these files are unique to FISHLIT. Contact details for first authors, in the form of postal addresses and/or URL or email addresses, are maintained up to date throughout the database to assist NISC disc users in requesting publications if the journals are unavailable to them.

The language or languages of both the article and the abstract are included as searchable fields, as is the article title.

The journal name, in the form of a coden, is entered along with the journal volume, journal number and publication year. The codens are standardised through the use of the authority list of codens produced by BIOSIS Serial Sources. Publications from almost three and a half thousand different sources are indexed for inclusion in FISHLIT and an important aspect of NISC indexing is to target grey literature as well as mainstream journal publications. The contributions by codens represented in BIOSIS is 46% whilst those included under "grey literature" make up the remaining 54%. These grey literature codens are assigned to publications that do not occur in the BIOSIS Serial Sources and include articles found in popular magazines and newspapers.

On average, 15 keywords are entered per record and these are also standardised, through the use of the ASFIS thesaurus. A high standard of indexing for FISHLIT is maintained by ensuring that all indexers are post-graduate students holding at least a Master of Science degree in one of the aquatic disciplines. This ensures that publications are scanned thoroughly and detailed keywords ascribed to each record.
The holding library for each reference is included in the database but is not as yet one of the searchable fields. This will change however in the near future, and once it is possible to search by this field, locating literature and document delivery will be considerably easier.

A particular strength of the FISHLIT database is the inclusion of the FISHLIT and Fisheries Review geolocators file. Indexers include as broad or as detailed geographic information as is provided in the article which significantly enhances the searchability of the records.

A second important feature of FISHLIT is the inclusion of taxonomic identifiers. As well as paying particular attention to including detailed geographic information, FISHLIT indexers take care to include taxonomic, both scientific and vernacular, identifiers if they are available in the text. The FISHLIT database is networked to the FISHNET database of the JLB Smith Institute fish collection thus access to constantly updated scientific

names is available. Family names are automatically inserted once the genus and species names have been entered into the database, eliminating any potential spelling errors.

All fields in the FISHLIT database undergo strict validation at several stages during the building of the database and records are thoroughly, manually edited before being included in the main FISHLIT file.

The current number of records in FISHLIT is approaching 83 000 and the database has rapidly grown from the small in-house system that it was, into an internationally recognised database. The recently expanded part-time staff body will ensure increased monthly inputs and extensive efforts are being made to expand the coverage of FISHLIT to include aquatic life forms other than fish, as well as being as comprehensive as possible in the fields of ichthyology, fisheries and aquaculture.

As mentioned earlier, FISHLIT is published on two NISC aquatic titles, Aquatic Biology, Aquaculture and Fisheries Resources, or ABAFR, and Fish and Fisheries Worldwide or FFW, and these two discs are now discussed in detail.

Fish and Fisheries Worldwide (FFW)
FISHLIT is published with nine other databases on Fish and Fisheries Worldwide, the largest being the Fisheries Review database that was formerly compiled by the US National Biological Service and is now incorporated in FISHLIT. Other databases included on FFW are Fish and Wildlife References Service, Fish Health News, Fish Viruses and Diseases, Castell's Nutrition References and AQUACULTURE. Figure 1 illustrates the contributions by each of these databases to the disc in terms of numbers of records.

FFW covers a wide variety of topics including aquaculture and fish diseases, ecology, physiology, fisheries, systematics, economics and pollution.

Two important databases have recently been included on this disc. Firstly, is a subset from the MEDLINE database dealing with fish diseases in aquaculture species, the effects of a fish diet on human health, fish oils, and fish and shellfish poisoning. Secondly is the database of the Fishing Industry Research Institute (FIRI) built by the CSIR in Cape Town, South Africa, which deals with aquaculture, fishing industries, and fish and shellfish processing and products.

Aquatic Biology, Aquaculture and Fisheries Resources (ABAFR)
The second disc hosting FISHLIT is Aquatic Biology, Aquaculture and Fisheries Resources which incorporates the whole of Fish and Fisheries Worldwide with Part 1, the living resources part, of the well-known ASFA database, and the relevant fisheries and aquaculture subsets of the CAB (Commonwealth Agricultural Bureau) Abstracts and AGRIS files (Figure 2).

These major collections under one CD-ROM title provide unparalleled access to information on the science and management of aquatic organisms and their environments

FFW and ABAFR are both available on CD-ROM and Internet.

Growth of FISHLIT

Figure 3 illustrates the growth of FISHLIT over the past decade by showing the number of new records added to FISHLIT expressed as a percentage of the total number of new records added to Fish and Fisheries Worldwide annually since 1985. There has been a gradual but steady increase in the contribution by FISHLIT records that is clearly illustrated by the percentage contribution. The significant increase in this contribution coincides with the hiring of a permanent indexer in 1995, as does the further increase in 1997.

The number of records added for 1998 is exceptionally low and it should be noted that this lag is expected for current year publications. Due to waiting for journals to be accessioned through the libraries, and because the information on the discs is updated once every three months, it is usual for additions of current year publications to only become obvious towards the end of the year and the beginning of the following year.

Figure 4 also serves to illustrate the growth of FISHLIT by comparing it to the growth of the well-known ASFA database by using a general search on fish including the keyterms "pisces" and any word beginning with "ichthy". As expected the number of new records added to ASFA is considerably higher than those added to FISHLIT. There are two reasons for this. Firstly, ASFA was initiated in the early 1960's whilst FISHLIT was only conceived in 1985. Secondly, ASFA has considerably more input centres around the world with a much larger staff base. However it is worth noting that there is a definite upward trend of new records being added to FISHLIT which will not only be maintained but also improved upon.

FISHLIT Source Material

Each publication included in the FISHLIT database is assigned a type code indicating the type of literature in which the publication is found. Figure 6 illustrates the proportions of types indexed for FISHLIT. The majority, almost 75%, is scientific articles from mainstream journals. Reports and popular articles are the next most important types and, to illustrate relative proportions, the contributions by short communications, books and conference proceedings were also plotted. There are 46 type descriptions currently in use and other examples include pamphlets, theses, bibliographies, atlases, workshops, microfiches, films, letters, appendices and annexes.

FISHLIT Subject coverage

Publications indexed for FISHLIT are assigned one or more subject codes broadly describing the main topics covered in each article. These are numerical codes describing 86 categories, and 8 examples of the more common ones used are given in Table 1. The marine, freshwater and estuarine categories each include fish, molluscs, crustaceans, algae, plankton, invertebrates, mammals and birds. These figures do not and should not add up to 100% as some publications are assigned categories such as catadromous or anadromous fish, civil engineering, climate and weather, and legal aspects and may not have also been given an aquatic environment type code.

Codes are also assigned to indicate whether publications include color or black and white illustrations, or checklists.

Although FISHLIT has mainly focused on fish, extensive efforts are now being made to expand coverage and include aspects of invertebrate biology and management as well as plankton and algae.

Table 1. Subject coverage for FISHLIT (September 1998).

Subject	Number	Percentage
Aquaculture	14 006	18
Fisheries	11 517	15
Management, education, policy, legislation	4 035	5
Physiology, biology, biochemistry	24 930	33
Marine	29 630	38
Freshwater	27 446	35
Estuarine	3 031	4
Ecology, conservation	14 017	18
Taxonomy	6 610	9

Bibliometric Analysis

Comparisons were drawn between FISHLIT and other databases by considering a number of example searches, all of which were done using the most recent (June 1998) release of the comprehensive anthology disc, Aquatic Biology, Aquaculture and Fisheries Review. Search examples were conducted under the broad topics of fisheries and aquaculture, as well as an example using scientific names. The keywords used and the number of hits for each search are given above the pie charts that illustrate the relative contribution to the search results by each of the databases in terms of numbers of records. Although ASFA has been accepted as the industry standard up until now, the following figures clearly illustrate the importance of integrating many databases on a single disc.

Fisheries (Figure 6)

FISHLIT contributed the most records, followed by ASFA and Fisheries Review, to the first search done on marine reserves in Africa. Almost equal contributions were made by FISHLIT, ASFA and Fisheries Review in the search using marine fishery management

and Africa as the keyterms. Although ASFA makes large contributions to the remaining two searches, one on artisanal and subsistence fishing and one on fishing gear, the benefit of including the other smaller, but nonetheless significant databases is clear as one is getting access to about fifty percent more information.

The composite record contribution is significant in all these searches, an important point considering that these are the super, value-added records created from records common to more than one database file.

Aquaculture (Figure 7)
The second set of search examples concerned aquaculture and introduces the newly added MEDLINE subset that contributed large proportions to the fish diseases and particularly the fish oils searches. Databases such as CAB and AGRIS featured more prominently under these aquaculture-orientated searches. FISHLIT played a rather significant role in the aquarium culture or ornamental fish search, and both ASFA and FISHLIT contributed significantly to the cage culture search.

Scientific names (Figure 8)
This set of searches done using scientific names illustrated the importance of taxonomic identifiers. As mentioned earlier, FISHLIT indexers take care to include as much taxonomic detail as possible and the FISHLIT database is networked to the FISHNET database of the JLB Smith Institute fish collection, allowing access to constantly updated scientific names.

ASFA, Fisheries Review and FISHLIT all contributed almost equally to the tilapia search results, whilst AGRIS and CAB made slightly smaller contributions. In the orange roughy search, ASFA and FISHLIT contributed the most whilst the abalone and catfish searches follow a similar pattern to the first two searches, with ASFA and FISHLIT contributing the largest number of records.

Conclusions
In conclusion, FISHLIT is a rapidly expanding aquatic science database aiming to comprehensively cover ichthyology, fisheries and aquaculture publications, with a particular emphasis on the African literature. By increasing the topic coverage of literature scanned, this database will become more meaningful to a broader range of aquatic scientists. Since no single database can provide 100% coverage, as illustrated through several search examples, the value of FISHLIT as well as that of the other databases with it is integrated, is increased through anthology publishing. Another important point illustrated from the search examples is the use of anthology discs in identifying and evaluating databases, an exercise especially useful to persons intending to purchase bibliographic information.

References

Crampton, M. 1988. FISHLIT: A Bibliographic Information System for Ichthyologists, Fisheries Scientists, and Fish Culturalists. In: *Marine Science Information Throughout the World.* (ed. by Winn, Burkhart & Burkhart). pp. 91-99. IAMSLIC.

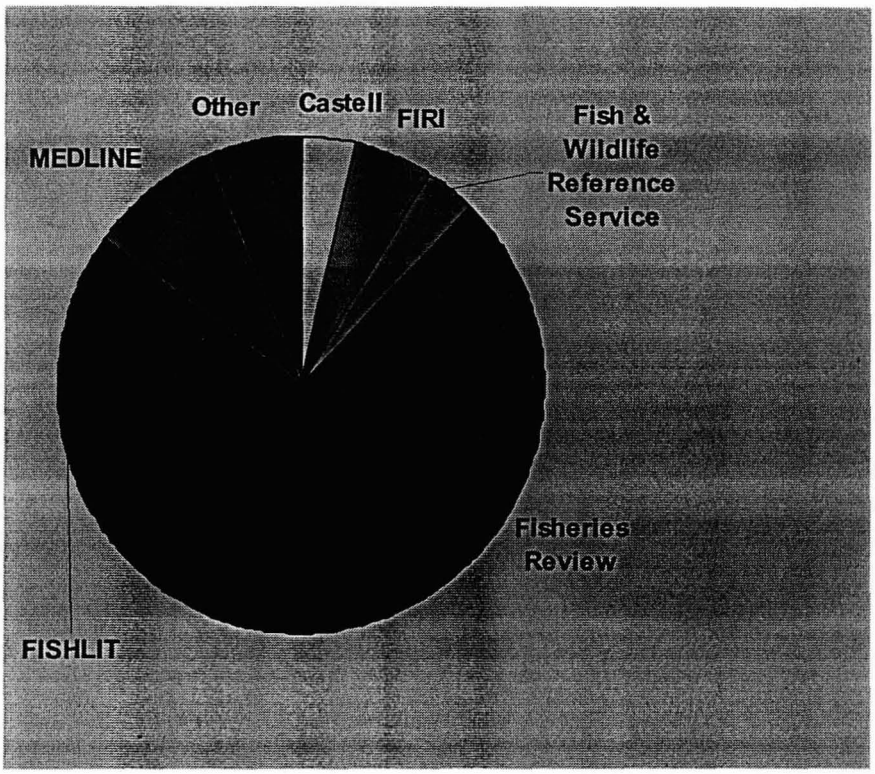

Figure 1. Pie chart illustrating make up of Fish and Fisheries Worldwide (August 1998) and relative contributions by each database in terms of numbers of records.

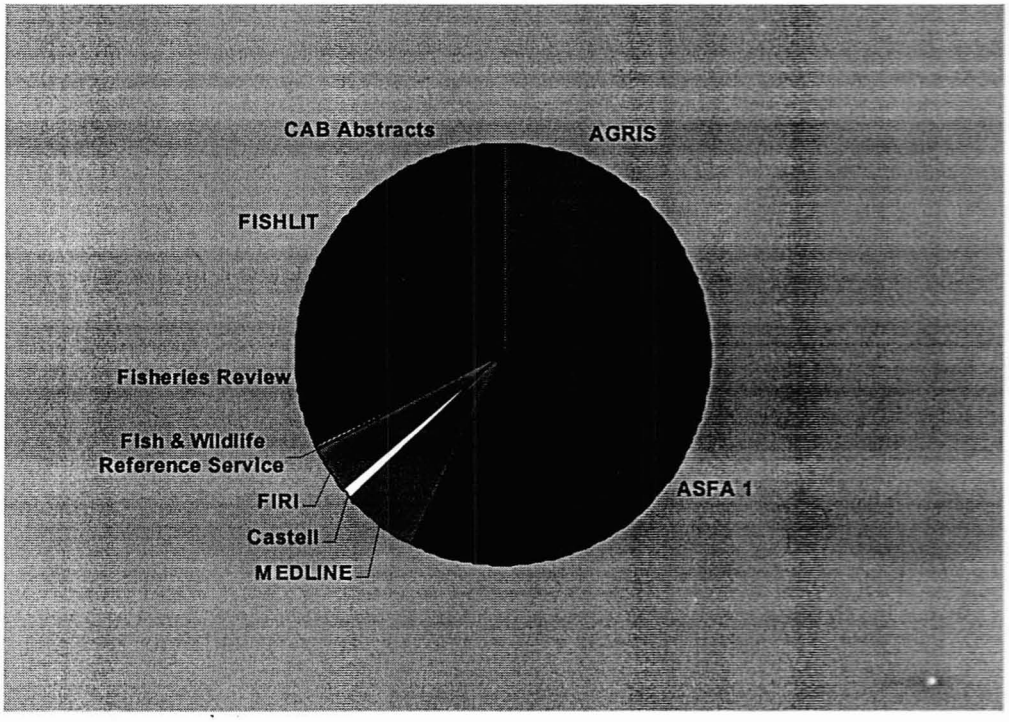

Figure 2. Pie chart illustrating make up of Aquatic Biology, Aquaculture and Fisheries Resources (June 1998) and relative contributions by each database.

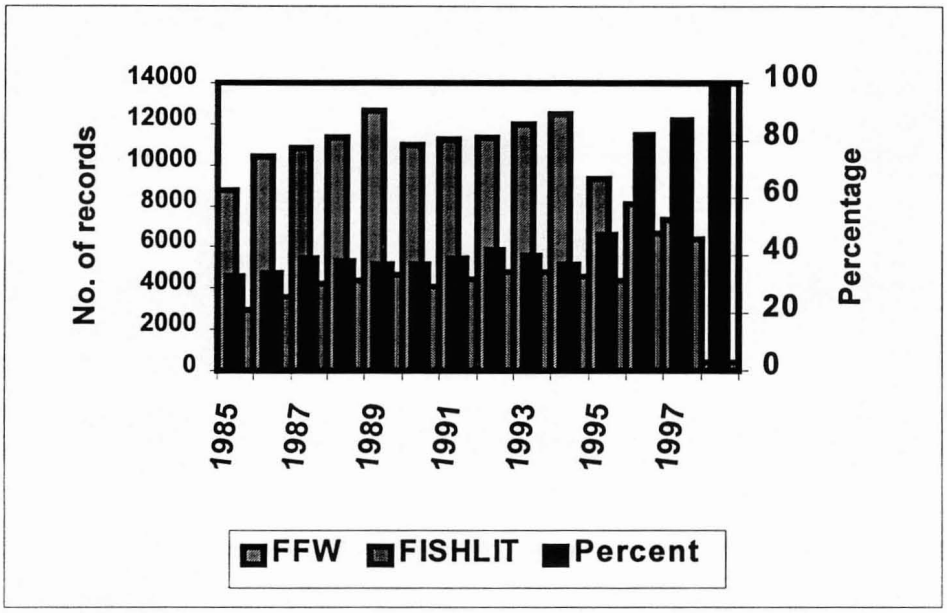

Figure 3. Growth of FISHLIT database illustrated by number and percentage contribution by new records added to Fish and Fisheries Worldwide annually since 1985.

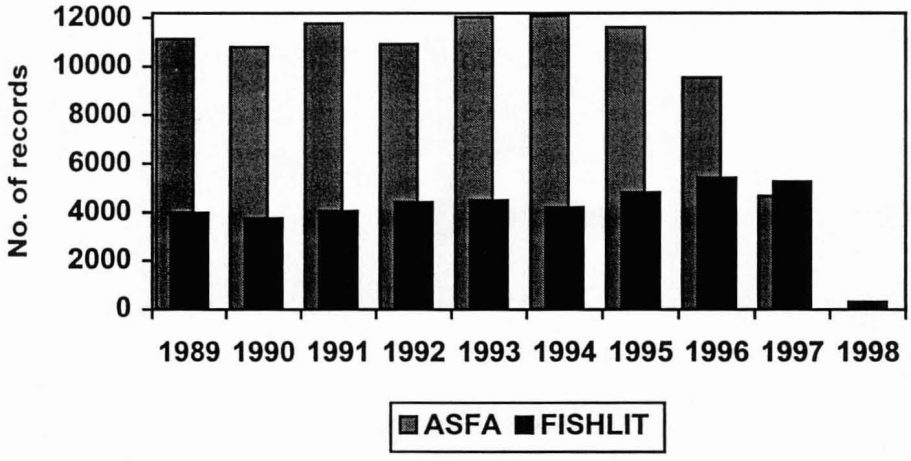

Figure 4. Growth of FISHLIT compared to growth of ASFA database demonstrated using the search "fish* or pisces or ichthy*" . Search was performed on Fish and Fisheries Worldwide (August 1998).

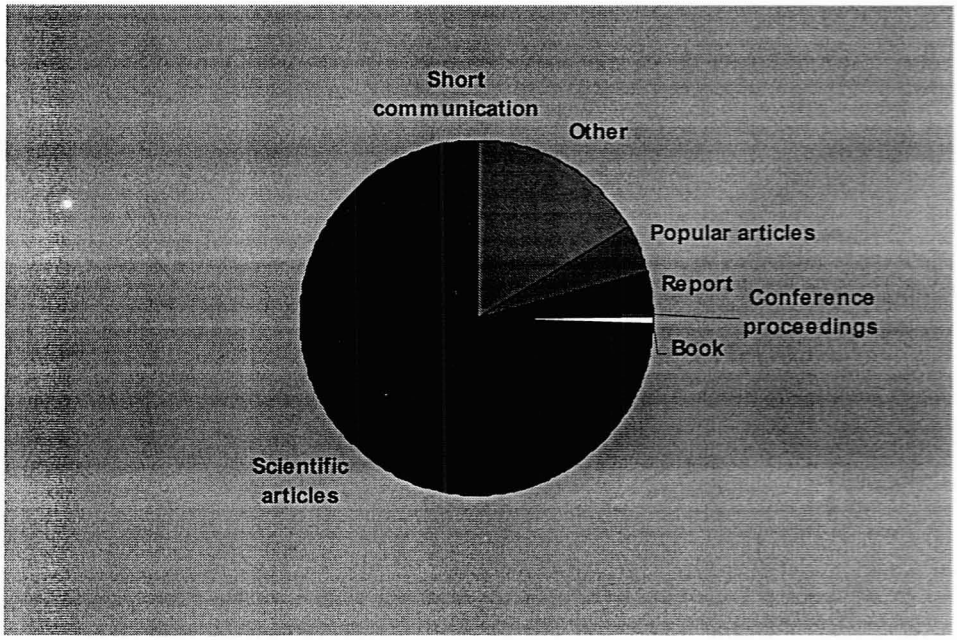

Figure 5. Pie chart illustrating percentage contributions by some of the publication types indexed for the FISHLIT database (September 1998).